Lecture Notes in Computer Science 800

Jan-Olof Eklundh (Ed.)

Computer Vision – ECCV '94

Third European Conference on Computer Vision
Stockholm, Sweden, May 1994
Proceedings, Volume I

Springer-Verlag

Jan-Olof Eklundh (Ed.)

Computer Vision – ECCV '94

Third European Conference on Computer Vision
Stockholm, Sweden, May 2-6, 1994
Proceedings, Volume I

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Recognition I

Shape Modelling

Shape Estimation

A Direct Recovery of Superquadric Models in Range Images Using Recover-and-Select Paradigm*

Aleš Leonardis, Franc Solina, and Alenka Macerl**

Computer Vision Laboratory, Faculty of Electrical Engineering and Computer Science
University of Ljubljana, Tržaška 25, 61001 Ljubljana, Slovenia
E-mail: Ales.Leonardis@fer.uni-lj.si, Franc.Solina@fer.uni-lj.si

Abstract. We present a novel approach to reliable and efficient recovery of part-descriptions from range images. We show that a set of superquadric models can be *directly* recovered from unsegmented range data, as opposed to methods which attempt the recovery of volumetric models only after the data has been pre-segmented using extensive pre-processing. The approach is based on the *recover-and-select* paradigm which consists of two intertwined stages: model-recovery and model-selection. At the model-recovery stage a redundant set of superquadrics is initiated in the image and allowed to grow, which involves an iterative procedure combining data classification and parameter estimation. All the recovered models are passed to the model-selection procedure where only the models resulting in the simplest overall description are selected.

1 Introduction and Motivation

The significance of detecting geometric structures in images has long been realized in the vision community. One of the primary intentions has been to build primitives that would bridge the gap between low-level features and high-level symbolic structures useful for further processing. Many theories have emerged which emphasize the importance of extraction of perceptually relevant image structures [14]. These structures essentially encode the knowledge or expectations of how the data is structured and help augmenting imperfect visual data with intrinsic information, thus enabling the recovery process to be more robust.

To represent the "natural" structuring of the world and support recognition of such "natural" structures from images, people employ a part structure. Perceptually, the world can be broken down into parts, and a goal of computer vision is to recover from images this part structure (segmentation) and the metric properties of individual parts (shape recovery). Two types of volumetric models, generalized cylinders and superquadrics, have emerged for such part-level modeling. Although rigorous schemes have been developed for the recovery

* This work was supported in part by The Ministry for Science and Technology of The Republic of Slovenia (Project P2-1122).

** Alenka Macerl is now with Hermes-Soft, d.o.o., Ljubljana, Slovenia.

Lecture Notes in Computer Science, Vol. 800
Jan-Olof Eklundh (Ed.)
Computer Vision – ECCV '94

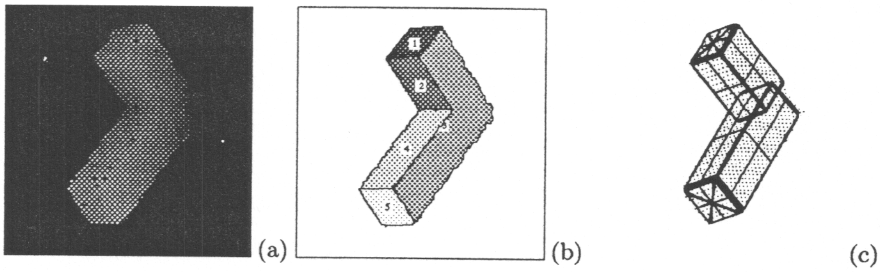


Fig. 1. (a) Range image, (b) surface-level description, (c) volumetric-level description.

of volumetric models, most of them make the assumption that the segmentation problem has been solved by some other means [19,17,21]. These approaches usually involve several steps, mostly applied in a hierarchical fashion, ranging from the estimation of local surface properties, curvature, etc., to more complex, such as symmetry seeking, in order to partition the data into parts that can supposedly be represented with a single volumetric model. These approaches, in fact, isolate the segmentation from the representation stage and significant efforts are necessary to combine, usually surface type descriptions, into volumetric models. The ability to even identify a set of surfaces as belonging to a given volume is not a trivial task without knowing at least connectedness and preferably surface closure. Moreover, a surface-level description may not be consistent with a volumetric description. Fig. 1 shows an L-shaped object whose volumetric description can not be obtained by a simple combination of recovered surfaces.

Pentland [14] proposed the use of superquadric models combined with global deformations as a set of primitives which could be recovered directly from images. The idea to use superquadrics for image interpretation has had quite a follow-up. Pentland [15,16,17], Solina and Bajcsy [19], Ferrie *et al.* [6], and Gupta and Bajcsy [8] used superquadrics in various ways to obtain part-level models. We think, however, that in all these approaches the predictive power of superquadrics for integral image interpretation (segmentation/shape recovery) has not been used to its full potential since other intermediary lower level models were used. We believe that when higher level generic models are well defined, as in the case of superquadrics, one can attempt to find them efficiently and reliably in a more direct way so that the models are employed to guide detection and grouping processes [4,2]. In this paper we demonstrate that superquadrics can be *directly* recovered from unsegmented range data. To achieve this goal we have cast the problem of volumetric recovery [19] in the *recover-and-select* paradigm for the recovery of geometric parametric structures from image data [10], which has already been successfully applied in segmentation and recovery of surface models in range data [12] and curve-models in intensity images [11].

The paper is organized as follows: section 2 is on *superquadric* models. In section 3 we outline the *recover-and-select* paradigm. Section 4 describes the details that pertain to the recovery of superquadric models. Some of our experimental results are shown in section 5.

2 Superquadric Models

The criteria for the selection of geometric primitives have been studied extensively by vision researchers [5]. An essential requirement is that the primitives are computable (*accessible*) from an input. Besides, they should balance the trade-off between data reduction and faithfulness to measured data. Superquadric models, which satisfy most of the criteria, are an extension of basic quadric surfaces and solids. They are used as volumetric primitives for shape representation in computer graphics [3], robotics [1], and computer vision [14,19,21,13,6] because they provide convenient part-level descriptions that can be further deformed and glued together to model articulated objects. However, as with any model-based approach, there are limitations regarding the expressive power of the models.

Superquadric surface is defined by the following equation

$$F(x,y,z) = \left(\left(\left(\frac{x}{a_1}\right)^{\frac{2}{\epsilon_2}} + \left(\frac{y}{a_2}\right)^{\frac{2}{\epsilon_2}} \right)^{\frac{\epsilon_2}{\epsilon_1}} + \left(\frac{z}{a_3}\right)^{\frac{2}{\epsilon_1}} \right) \quad ; \ \epsilon_1, \epsilon_2 > 0 \ . \qquad (1)$$

When both exponents ϵ_1 and ϵ_2 equal 1, the surface is an ellipsoid. When $\epsilon_1 \ll 1$ and $\epsilon_2 = 1$, the superquadric surface is shaped like a cylinder. Parallelepipeds are produced when both $\epsilon_1 \ll 1$ and $\epsilon_2 \ll 1$. Modeling capabilities of superquadrics can be enhanced by deforming them in different ways using global [19] or local deformations [21].

2.1 Recovery of Superquadrics

Here we summarize a robust method for the recovery of isolated superquadric models [19]. The method, which is based on a nonlinear least-squares method, has been successfully used by several researchers [1,8,18,6].

The implicit function (Eq. 1) defined in an object centered coordinate system determines where a given point (x, y, z) lies relative to the superquadric surface. To recover a superquadric in general position, the implicit function for general position is used

$$F(x,y,z) = F(x,y,z;a_1,a_2,a_3,\epsilon_1,\epsilon_2,\phi,\theta,\psi,p_x,p_y,p_z) \ . \qquad (2)$$

This expanded "inside-outside" function has 11 parameters; a_1, a_2, a_3 define the superquadric size; ϵ_1 and ϵ_2 are shape parameters; ϕ, θ, ψ define the orientation, and p_x, p_y, p_z define the position in space.

Suppose we have a set of 3-D points on a surface of an object which we want to model with a superquadric. Using a non-linear iterative least squares method (Levenberg-Marquardt) we can find such a set of model parameters that most of the 3-D points will lie on, or close to the model's surface. Since due to self occlusion not all sides of an object are visible at the same time, we search for the *smallest* superquadric that fits the given range points in the least squares sense by expanding the fitting function in the following way:

$$f = \sqrt{a_1 a_2 a_3}\,(F^{\epsilon_1} - 1) \ . \qquad (3)$$

To make the error metric more quadratic and more suited for rapid convergence, we raised the "inside-outside" function F (Eq. 1) to the power of ϵ_1 in Eq. (3). Since only very rough initial estimates of object's true position, orientation, and size suffice to assure convergence to a local minimum that corresponds to the actual shape, the initial values for ϵ_1 and ϵ_2 are set to 1. Position is estimated by computing the center of gravity of all range points, and the orientation by computing the central moments with respect to the center of gravity. To prevent the minimization procedure from getting stuck in a shallow local minimum, we add noise during the minimization procedure.

3 Recover-and-Select Paradigm

We present here a brief general outline of the recover-and-select paradigm. For details the reader is referred to [10]. The choice of specific models, in our case superquadrics, imposes certain constraints on the recovery of these structures from images, which will be described in section 4.

3.1 Model Recovery

Recovery of parametric models from unsegmented data is difficult because one has to solve two interrelated problems: *find image elements that belong to a single parametric model* **and** *determine the values of the parameters of the model.*

For image elements that have already been classified (segmented) one can determine the parameters of a model by applying standard statistical estimation techniques. Conversely, knowing the parameters of the model, a search for compatible image points can be accomplished by pattern classification methods. Thus we propose to solve these **two** problems simultaneously by an iterative method, conceptually similar to the one described by Besl [4], which combines data classification and model fitting.

One of the crucial dilemmas is where to find the initial estimates (seeds) in an image since their selection has a major effect on the success or failure of the overall procedure. We propose that a search for the points that could belong to a single parametric model is performed in a grid-like pattern of windows overlaid on the image. Thus, the requirement of classifying all data points of a certain model is relaxed to finding only a small subset. However, there is no guarantee that every seed will lead to a good description since some initial models can be constructed over areas which belong to different models. As a remedy we propose to *independently* build *all possible* models using all statistically consistent seeds and to use them as hypotheses that could compose the final description.

Having an initial set of points (a seed) we estimate the parameters of the model. If sufficient similarity between the model and the data is established, ultimately depending on the task at hand, we proceed with a search for more compatible points. An efficient search is performed in the vicinity of the current border points of the model. New compatible image elements are included in the data set and the parameters of the model are updated. The new goodness-of-fit is computed and compared to the old value. This is followed by a decision whether to perform another iteration or terminate the procedure.

The final outcome of the model-recovery procedure for a particular model M_i consists of three terms which are subsequently passed to the model-selection procedure: the set of data elements that belong to the model M_i, the corresponding set of parameters of the model, and the goodness-of-fit value which describes the conformity between the data and the model.

The main features of the model-recovery procedure are: a high degree of resistance to outliers since the performance of the fitting is constantly monitored and an independent and potentially parallel execution of the recovery procedure for individual models.

3.2 Model Selection

The redundant representation obtained by the model-recovery procedure is a direct consequence of the decision that a search for parametric volumetric models is initiated everywhere in an image. Several of the models are completely or partially overlapped. The task of combining different models is reduced to a selection procedure which selects those models that produce the simplest description, i.e., the one that describes the data with the minimum number of models while keeping the deviations between data points and models low. Intuitively, this reduction in complexity of a representation coincides with a general notion of simplicity which has a long history in psychology (Gestalt principles). The formalization of this principle led in information theory to the method of *Minimum Description Length* MDL, which has recently found its way to computer science, including computer vision [9,7,16].

The objective function $F(\mathbf{m})$, which is to be maximized in order to produce the "best" description in terms of models, has the following form:

$$F(\mathbf{m}) = \mathbf{m}^T \mathbf{Q} \mathbf{m} = \mathbf{m}^T \begin{bmatrix} c_{11} & \dots & c_{1M} \\ \vdots & & \vdots \\ c_{M1} & \dots & c_{MM} \end{bmatrix} \mathbf{m} \ , \tag{4}$$

where $\mathbf{m}^T = [m_1, m_2, \dots, m_M]$. m_i is a *presence variable* having the value 1 for the presence and 0 for the absence of the model M_i in the final description. The diagonal terms of the matrix $\mathbf{Q}$ express the cost-benefit value for a particular model M_i. This value is a function of the number of data points that belong to the model, the goodness-of-fit measure between the model and the data, and the cost of specifying the parameters of the model. The off-diagonal terms handle the interactions between the overlapping models, taking into account the mutual error and the number of data points covered by both models.

Maximizing the objective function $F(\mathbf{m})$ belongs to the class of problems known as combinatorial optimization (Quadratic Boolean problem). Since the number of possible solutions increases exponentially with the size of the problem, it is usually not tractable to explore them exhaustively. Thus the exact solution has to be sacrificed to obtain a practical one. It turned out that in our case, for well-behaved inputs (in the sense of being well describable by the chosen set of models), we obtain reasonable solutions by a direct application of the *greedy algorithm* which at any individual stage selects the option which is locally optimal. In other words, the models are selected in the sequence that corresponds to the size of their contributions to the objective function, which is equivalent to applying at each stage of the algorithm the *winner takes all* principle.

3.3 Model Recovery and Model Selection

In order to achieve a computationally efficient procedure the model-recovery and model-selection procedures are combined in an iterative fashion. The recovery of currently active models is interrupted by the model-selection procedure which selects a set of currently optimal models which are then passed back to the model-recovery procedure. This process is repeated until the remaining models are completely recovered. The trade-offs which are involved in the dynamic combination of these two procedures are discussed elsewhere [10].

4 Superquadrics in the Recover-and-Select Paradigm

In this section we give some details of the superquadric model recovery as used in the recover-and-select paradigm. The overall processing scheme is summarized in Table 1. The seed selection, the goodness-of-fit measure used in decision making, and the search for new compatible points (growing) are briefly explained in the following subsections.

4.1 Seed Selection

Initial models (seeds) are placed on the image in a grid-like pattern of windows. An initial seed encompasses a set of range data points in a small window whose size is determined on the basis of scale and can be adaptively changed depending on the task. A superquadric model is then fitted to each of the data sets. Next, a decision is made whether a seed is allowed to grow further or it is rejected. This decision, which is based on the goodness-of-fit measure (see the next subsection), helps eliminating the seeds contaminated with outliers and those seeds that were placed on data sets that cross part boundaries. Thus the number of initial models at the start of processing is reduced.

4.2 Goodness-of-Fit Measure

A decision whether a model should grow further or not depends on the established similarity between the model and the data. If sufficient similarity is established, we accept the currently estimated parameters, together with the current data set, and proceed with the search for more compatible points. The question is what could be used as a goodness-of-fit measure. Due to its dependence

```
input: a range image

determine a set of seeds

for all seeds do
    fit a SQ model (estimate parameters)
    if (GOF == OK) put the SQ model into the set of currently active, not fully grown
models
    else the seed is rejected

while there are any active, not fully grown SQ models do
    for all active not fully grown SQ models do
        extrapolate/find compatible points
        if no new compatible points
            the SQ model is fully grown
        else
            fit a SQ model
            if (GOF ≠ OK)
                reject the last included points
                the SQ model is fully grown
            end if

        end if
    end for

    perform selection among all active SQ models for the current optimal description
        (only the selected SQ models remain active)

end do

output: part-level (SQ) description of a range image
```

Table 1. Superquadric models in the *Recover-and-Select* paradigm.

on the superquadric size and shape parameters $(a_1, a_2, a_3, \epsilon_1, \epsilon_2)$ the algebraic distance (inside-outside function (1)) is not suitable. An approximation to the Euclidean distance does the job when the distance of a point from the corresponding superquadric model is small. In fact, due to the strategy of acquiring new data points based on the consistency with the current model, this condition is always satisfied. The approximate Euclidean distance for a given data point is:

$$d^2(x,y,z) \approx \frac{f^2(x,y,z)}{\| \nabla f((x,y,z)) \|^2} , \tag{5}$$

where f is given in equation (3). The sum over all data points belonging to the model determines the goodness-of-fit of the entire model. This is also the measure which is, together with the number of data points encompassed by the model, passed to the selection procedure.

4.3 Search for New Compatible Points

In accordance with the paradigm, an efficient search for new compatible points is performed in the vicinity of the present border points of the model. This is achieved first by simply increasing the values of the parameters (a_1, a_2, a_3) of the current model to get an enlarged model. Then all the points which are inside this enlarged model are checked on how close they are to the surface of the original model (again, the approximate Euclidean distance (5) is used). Only those points that are close enough to the original model are included in the updated set of points. On this set of points a new superquadric model-recovery procedure is started.

5 Experimental Results

The proposed method has been tested on a variety of range images. The original range images were subsampled to speedup the recovery. Processing of the examples shown took on a workstation HP-715/50 on the average less than 10 minutes. However, processing time can be significantly reduced on a parallel machine since individual models can be recovered in parallel.

Three examples of processed range images are shown[3]. Each figure shows the following image sequence: **(a)**—the intensity image of the object, **(b)**—the corresponding range image, **(c)**—the original image resampled and transformed into the form (x,y,z), appropriate for the superquadrics recovery process, **(d)**—the recovered volumetric models after the first model selection, **(e)**—volumetric models at the midpoint of the recover-and-select process, **(f)**—final result (superquadric models on top of only those range points that influenced the models). The results are commented in the captions of each figure.

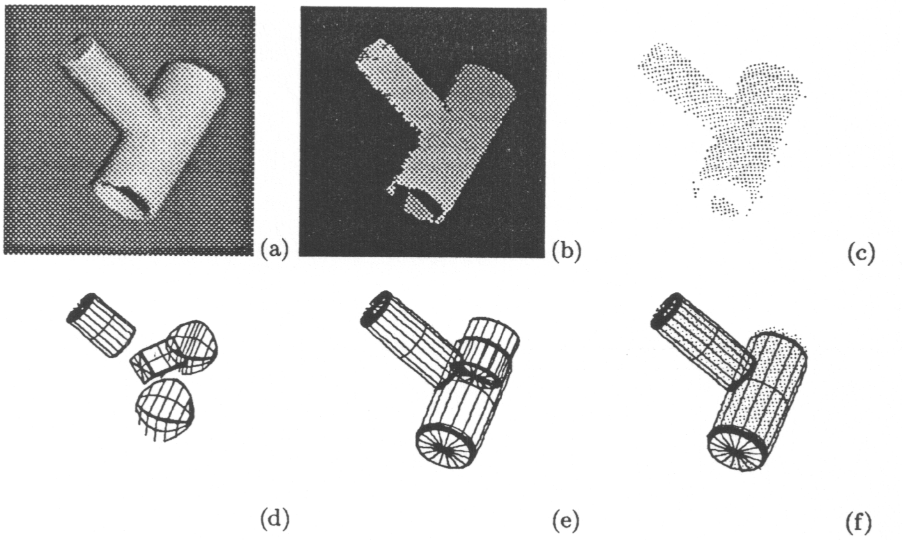
6 Conclusions

We have successfully combined two existing methods, namely recovery of superquadric models [19] and the *recover-and-select* paradigm [10]. Thus we showed that a direct segmentation into part-level volumetric models is possible. The interpretation of these models is straightforward, with direct applications for manipulation, object recognition, and CAD modeling (reverse engineering).

There are several open issues that we are going to address in our future work. An immediate problem is merging of the models that represent different sides of the same physical part separated by occlusion (see Fig. 3).

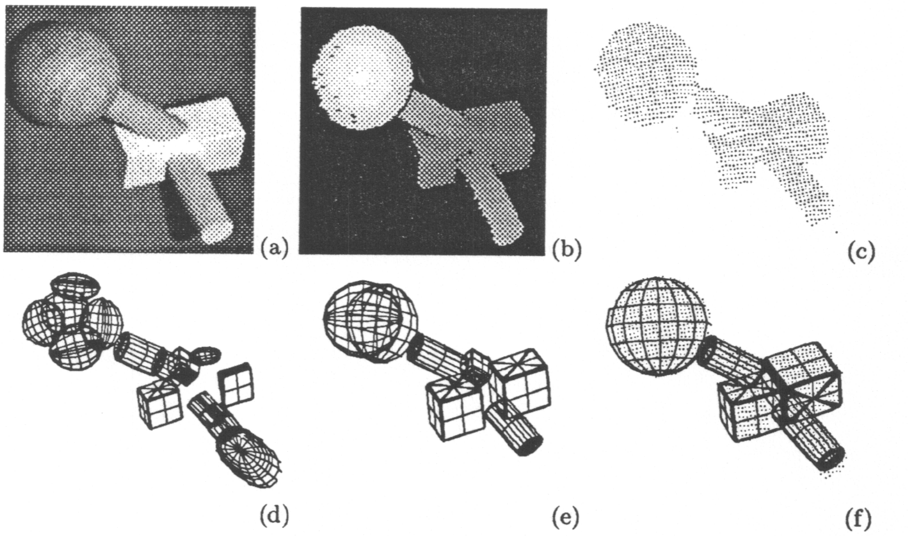
We also plan to explore the possibility to use the proposed method for recovering (in parallel) multiple geometric modalities (surfaces and volumes) and different image modalities (range and intensity images). The final description of a scene would result from a selection procedure (employing the MDL principle), as proposed in [20].

[3] All range images shown in this paper were kindly provided by Marjan Trobina from ETH, Zürich, Switzerland.



(a) (b) (c) (d) (e) (f)

Fig. 2. A tube and a cylinder that intersect. Twenty seeds were placed on the range image at the start of processing. The final representation is most compact, consisting of two superquadrics. The points on the inside of the tube are also incorporated into the final description (f).



(a) (b) (c) (d) (e) (f)

Fig. 3. A complex scene consisting of a sphere and two cylinders that are joined with a block. Forty seeds were placed on the scene at start. The final description (f), however, consists of a sphere, two cylinders, and *two* blocks. Neither of the two superquadric models in figure (e) could grow into the other side of the block because the number of range points that bridge the gap caused by occlusion is not sufficient.

References

1. P. K. Allen and P. Michelman. Acquisition and interpretation of 3-D sensor data from touch. *IEEE Trans. on Robotics and Autom.*, 6(4):397–404, 1990.
2. R. Bajcsy, F. Solina, and A. Gupta. Segmentation versus object representation—are they separable? In R. Jain and A. Jain, editors, *Analysis and Interpretation of Range Images*, Springer-Verlag, New York, 1990.
3. A. H. Barr. Superquadrics and angle-preserving transformations. *IEEE Comp. Graph. and Appl.*, 1(1):11–23, 1981.
4. P. J. Besl. *Surfaces in Range Image Understanding.* Springer-Verlag, 1988.
5. M. Brady, J. Ponce, A. Yuille, and H. Asada. Describing surfaces. *Computer Vision, Graphics, and Image Processing,* 32(1):1–28, 1985.
6. F. P. Ferrie, J. Lagarde, and P. Whaite. Darboux frames, snakes, and superquadrics: Geometry from the bottom up. *IEEE Trans. on* PAMI-15(8):771–784, 1993.
7. P. Fua and A. J. Hanson. Objective functions for feature discrimination. In *Proc. 11th Int. Joint Conf. on Artif. Intell.*, pp. 1596–1602, Detroit, MI, 1989.
8. A. Gupta and R. Bajcsy. Volumetric segmentation of range images of 3D objects using superquadric models. *Computer Vision, Graphics, and Image Processing: Image Understanding,* 58(3):302–326, 1993.
9. Y. G. Leclerc. Constructing simple stable descriptions for image partitioning. *Int. Journal of Computer Vision*, 3:73–102, 1989.
10. A. Leonardis. *Image Analysis Using Parametric Models: Model-Recovery and Model-Selection Paradigm.* A Dissertation in Computer Science. University of Ljubljana, Faculty of EE & CS. May 1993.
11. A. Leonardis and R. Bajcsy. Finding parametric curves in an image. In G. Sandini, editor, *Proc. of The Second European Conf. on Computer Vision—ECCV-92*, pp. 653–657. Springer-Verlag, 1992. LNCS-Series Vol. 588.
12. A. Leonardis, A. Gupta, and R. Bajcsy. Segmentation of range images as the search for geometric parametric models. *Int. Journal of Computer Vision*, To appear.
13. D. Metaxas and D. Terzopoulos. Shape and nonrigid motion estimation through physics-based synthesis. *IEEE Trans. on* PAMI 15(6):580–591, 1993.
14. A. P. Pentland. Perceptual organization and the representation of natural form. *Artif. Intell.*, 28(2):293–331, 1986.
15. A. P. Pentland. Recognition by parts. In *Proc. First Int. Conf. on Computer Vision*, London, pp. 612–620, 1987.
16. A. P. Pentland. Automatic extraction of deformable part models. *Int. Journal of Computer Vision,* 4:107–126, 1990.
17. A. P. Pentland and S. Sclaroff. Closed-form solutions for physically based shape modeling and recognition. *IEEE Trans. on* PAMI 13(7):715–729, 1991.
18. N. S. Raja and A. K. Jain. Recognizing geons from superquadrics fitted to range data. *Image and Vision Computing Journal,* 10(3):179–190, 1992.
19. F. Solina and R. Bajcsy. Recovery of parametric models from range images: The case for superquadrics with global deformations. *IEEE Trans. on* PAMI-12(2):131–147, 1990.
20. F. Solina and A. Leonardis. Selective scene modeling. In *Proc. of the 11th Int. Conf. on Pattern Recognition*, pp. A:87–90, The Hague, The Netherlands, 1992.
21. D. Terzopoulos and D. Metaxas. Dynamic 3D models with local and global deformations: Deformable superquadrics. *IEEE Trans. on* PAMI 13(7):703–714, 1991.